

#28
7-25-01

OIPE

RAW SEQUENCE LISTING
 PATENT APPLICATION: US/08/971,172A

DATE: 07/25/2001
 TIME: 15:27:30

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\07252001\H971172A.raw

SEQUENCE LISTING

ENTERED

- 2 (1) GENERAL INFORMATION:
 3 (i) APPLICANT: Goodman, Corey S.
 4 Kidd, Thomas
 5 Mitchell, Kevin
 6 Tear, Guy
 7 (ii) TITLE OF INVENTION: Robo: A Novel Family of Polypeptide and
 8 Nucleic Acids
 9 (iii) NUMBER OF SEQUENCES: 13
 10 (iv) CORRESPONDENCE ADDRESS:
 11 (A) ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP
 12 (B) STREET: 75 DENISE DRIVE
 13 (C) CITY: HILLSBOROUGH
 14 (D) STATE: CALIFORNIA
 15 (E) COUNTRY: USA
 16 (F) ZIP: 94010
 17 (v) COMPUTER READABLE FORM:
 18 (A) MEDIUM TYPE: Floppy disk
 19 (B) COMPUTER: IBM PC compatible
 20 (C) OPERATING SYSTEM: PC-DOS/MS-DOS
 21 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30
 22 (vi) CURRENT APPLICATION DATA:
 23 (A) APPLICATION NUMBER: US/08/971,172A
 24 (B) FILING DATE: 14-Nov-1997
 25 (C) CLASSIFICATION:
 26 (viii) ATTORNEY/AGENT INFORMATION:
 27 (A) NAME: OSMAN, RICHARD A
 28 (B) REGISTRATION NUMBER: 36,627
 29 (C) REFERENCE/DOCKET NUMBER: B98-006
 30 (ix) TELECOMMUNICATION INFORMATION:
 31 (A) TELEPHONE: (650) 343-4341
 32 (B) TELEFAX: (650) 343-4342
 34 (2) INFORMATION FOR SEQ ID NO: 1:
 35 (i) SEQUENCE CHARACTERISTICS:
 36 (A) LENGTH: 4188 base pairs
 37 (B) TYPE: nucleic acid
 38 (C) STRANDEDNESS: double
 39 (D) TOPOLOGY: linear
 40 (ii) MOLECULE TYPE: cDNA
 41 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
 42 ATGCATCCCA TGCATCCCGA AAACCACGCC ATCGCCCGGA GCACGAGCAC CACTAATAAC 60
 43 CCATCTCGCA GTCGGAGCAG CAGGATGTGG CTCCTGCCG CCTGGCTGCT CCTCGTCCTG 120
 44 GTGGCCAGCA ATGGCCCTGCC AGCAGTCAGA GGCCAGTACC AATGCCACG TATCATCGAG 180
 45 CATCCCACGG ATCTGGTCGT TAAGAAGAAT GAACCCGCCA CGCTCAACTG CAAAGTGGAG 240
 46 GGCAGGCCGG AACCCACCAT TGAGTGGTTT AAGGATGGCG AACCCGTCAAG CACCAACGAA 300
 47 AAGAAATCGC ACCGCGTCCA GTTCAAGGAC GGCGCCCTCT TCTTTTACAG GACAATGCAA 360
 48 GGCAAGAAGG AGCAGGACGG CGGAGAGTAC TGGTGCCTGG CCAAGAACCG AGTGGGCCAG 420

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/971,172A

DATE: 07/25/2001

TIME: 15:27:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07252001\H971172A.raw

49	GCCGTTAGTC	GCCATGCCTC	CCTCCAGATA	GCTGTTTGC	GCGACGATT	T CGCGTGGAG	480
50	CCCAAAGACA	CGCGAGTGGC	CAAAGGCGAG	ACGGCTCTGC	TGGAGTGTGG	GCCGCCAAA	540
51	GGCATTCCAG	AGCCAACGCT	GATTTGGATA	AAGGACGGCG	TTCCCTTGGGA	CGACCTGAAA	600
52	GCCATGTCGT	TTGGCGCCAG	CTCCCGCGTT	CGAATTGTGG	ACGGTGGCAA	CCTGCTGATC	660
53	AGCAATGTGG	AGCCCATTGA	TGAGGGCAAC	TACAAGTGCA	TTGCCAGAA	TCTGGTAGGC	720
54	ACCCGCGAGA	GCAGCTATGC	CAAGCTGATT	GTCCAGGTCA	AACCATACTT	TATGAAGGAG	780
55	CCCAAGGATC	AGGTGATGCT	CTACGGCCAG	ACAGCCACTT	TCCACTGCTC	AGTGGGCGGT	840
56	GATCCGCCGC	CGAAAGTGT	GTGGAAAAAG	GAGGAGGGCA	ATATTCCGGT	GTCCAGAGCG	900
57	CGAATCCTTC	ACGACGAGAA	AAGTTAGAG	ATATCCAACA	TAACGCCAAC	CGATGAGGGC	960
58	ACCTATGTCT	GCGAGGCACA	CAACAATGTC	GGTCAGATCA	GCGCTAGGGC	TTCTCTTATA	1020
59	GTCCACGCTC	CGCCGAACCT	TACGAAAAGA	CCCAGTAACA	AGAAAGTGGG	ACTAAATGGG	1080
60	GTTGTCCAAC	TACCTTGCAT	GGCCTCCGGA	AACCCTCCGC	CGTCTGTATT	CTGGACCAAG	1140
61	GAAGGAGTAT	CCACTCTTAT	GTTCCTAAAT	AGTTCGCACG	GAAGGCAGTA	TGTGGCTGCC	1200
62	GATGGAACTC	TGCAGATTAC	GGATGTGCGG	CAGGAAGACG	AAGGCTACTA	TGTGTGTTCC	1260
63	GCTTCAGTG	TAGTCGATT	CTCTACAGTA	CGGGTTTCC	TGCAAGTCAG	CTCGGTAGAC	1320
64	GAGCGTCCAC	CTCCGATTAT	TCAAATCGGA	CCTGCCAATC	AAACACTGCC	CAAGGGATCA	1380
65	GTTGCTACTT	TACCCTGTCG	GGCCACTGGG	AATCCCAGTC	CCCGTATCAA	GTGGTTCCAC	1440
66	GATGGACATG	CCGTACAAGC	GGGCAATCGA	TACAGCATCA	TCCAAGGAAG	CTCACTGAGA	1500
67	GTCGATGACC	TTCAACTAAG	TGACTCTGGT	ACCTACACCT	GCACTGCATC	TGGCGAACGA	1560
68	GGAGAAACCTT	CCTGGGCTGC	CACACTAACG	GTGGAAAAAC	CCGGTTCTAC	ATCTCTTCAC	1620
69	CGGGCAGCTG	ATCCTAGCAC	TTATCCTGCT	CCTCCAGGAA	CACCTAAAGT	CCTGAATGTC	1680
70	AGTCGCACCA	GCATTAGTCT	TCGTTGGGCT	AAAAGCCAAG	AGAAACCCGG	AGCTGTGGC	1740
71	CCAATCATTG	GATACACTGT	AGAGTACTTC	AGTCCGGATC	TGCAAACCTGG	TTGGATTGTG	1800
72	GCTGCCATC	GAGTCGGCGA	CACTCAAGTC	ACTATCTCGG	GTCTCACTCC	TGGCACTTCG	1860
73	TATGTGTTCC	TAGTTAGAGC	TGAGAATACT	CAGGGTATTT	CTGTGCCTTC	CGGCTTATCA	1920
74	AATGTTATTA	AAACCATTGA	GGCAGATTTC	GATGCAGCTT	CTGCCAATGA	TTTGTCAAGCA	1980
75	GCTCGAACCTT	TGCTGACAGG	AAAGTCGGTG	GAGCTAATAG	ATGCCTCGGC	TATCAATGCT	2040
76	AGTGCCTGTTA	GACTTGAGTG	GATGCTCCAC	GTGAGCGCTG	ATGAGAAAATA	CGTAGAGGGC	2100
77	CTGCGCATAC	ACTATAAGGA	TGCCAGTGT	CCATCCGCAC	AGTATCACTC	GATCACTGTT	2160
78	ATGGATGCCT	CTGCAGAACATC	GTTCGTGGTG	GGAAACCTTA	AGAAGTACAC	CAAGTATGAG	2220
79	TTCTTCCTAA	CACCCTTTT	TGAGACAATT	GAAGGACAGC	CCAGTAACTC	CAAGACAGCC	2280
80	CTCACCTATG	AAGATGTTCC	CTCCGCACCA	CCGGATAACAA	TTCAGATTGG	CATGTACAAC	2340
81	CAAACAGCCG	GTTGGGTCG	TTGGACTCCG	CCACCCCTCCC	AGCACCACAA	TGGCAATTG	2400
82	TATGGCTACA	AGATTGAGGT	CAGCGCCGGT	AAACACCATGA	AGGTGCTGGC	CAATATGACT	2460
83	CTTAATGCTA	CCACCCACATC	TGTGCTCCTA	AATAAACCTAA	CCACCGGAGC	TGTGTACAGC	2520
84	GTGAGGTTGA	ACTCCTTTAC	CAAGGCAGGA	GATGGACCTT	ACTCCAAACCC	GATATCACTA	2580
85	TTCATGGACCC	CCACCCATCA	TGTGCATCCG	CCACGGGCAC	ATCCAAGCGG	CACCCATGAT	2640
86	GGGCGACATG	AGGGACAGGA	TCTCACGTAT	CATAACAAATG	GCAACATACC	ACCTGGCGAC	2700
87	ATTAATCCCA	CCACTCATAA	AAAGACCACT	GACTACCTAT	CTGGACCGTG	GCTAATGGTG	2760
88	CTGGTCTGCA	TCGTTCTTCT	AGTCCTGGT	ATTCGGCGG	CTATTTGAT	GGTCTACTTC	2820
89	AAGCGCAAGC	ATCAAATGAC	CAAGGAATTG	GGTCACTTAA	GTGTGGTCAG	TGACAACGAA	2880
90	ATAACCGCAT	TAAATATCAA	TAGCAAAGAG	AGCCTTGGA	TAGACCATCA	TCGTGGATGG	2940
91	CGAACTGCCG	ATACTGACAA	AGACTCAGGA	TTAACGCGAAT	CGAAGCTACT	ATCCCACGTT	3000
92	AACAGCAGTC	AATCCAACTA	CAATAACTCC	GATGGAGGAA	CCGATTATGC	AGAAGTTGAC	3060
93	ACCCGTAACC	TTACCACCTT	CTACAATTGT	CGCAAGAGCC	CCGATAATCC	CACGCCGTAC	3120
94	GCCACCACTA	TGATCATTGG	TACCTCTTCC	AGTGAGACCT	GCACCAAGAC	AACATCTATA	3180
95	AGTGCCTGATA	AGGACTCGGG	AACTCATTG	CCCTATTCTG	ACGCATTG	CGGTCAAGGTG	3240
96	CCAGCGGTT	CTGTTGTCAA	ATCCAACAT	CTTCAGTATC	CGGTTGAACC	GATCAACTGG	3300
97	TCAGAGTTTC	TACCCCCGCC	GCCAGAACAC	CCACCTCCGT	CTTCTACCTA	TGGATACGCA	3360

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/971,172A

DATE: 07/25/2001

TIME: 15:27:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07252001\H971172A.raw

98	CAAGGATCTC	CTGAATCTTC	GCGGAAGAGC	TCCAAAAGCG	CAGGTTCCGG	CATTCTACA	3420
99	AATCAAAGCA	TTCTGAACGC	ATCCATACAC	AGCAGCTCCT	CGGGCGGCC	TTCAGCTTGG	3480
100	GGAGTATCGC	CCCAATATGC	TGTCGCCTGT	CCACCGGAAA	ACGTTTATAG	CAATCCGCTG	3540
101	TCGGCAGTGG	CTGGCGGCAC	CCAGAACCGC	TATCAGATAA	CGCCCACAAA	CCAACATCCG	3600
102	CCACAGTTAC	CGGCCTACTT	TGCCACCACG	GGTCCAGGAG	GAGCTGTACC	ACCCAACCAC	3660
103	CTGCCATTG	CCACACAGCG	TCATGCAGCC	AGCGAGTACC	AGGCTGGACT	GAATGCAGCG	3720
104	CGATGTGCC	AAAGCCGCGC	CTGCAACAGC	TGCGATGCCT	TGGCACACC	CTCGCCCATG	3780
105	CAACCCCCAC	CGCCAGTTCC	CGTACCCGAG	GGCTGGTACC	AACCGGTGCA	TCCCAATAGC	3840
106	CACCCGATGC	ACCCGACCTC	CTCCAACAC	CAGATCTACC	AGTGCTCCTC	CGAGTGCTCG	3900
107	GATCACTCGA	GGAGCTCGCA	GAGTCACAAG	CGGCAGCTGC	AGCTCGAGGA	GCACGGCAGC	3960
108	AGTGCCAAAC	AACGCCGGAGG	ACACCACCGT	CGACGAGCCC	CGGTGGTGC	GCCGTGCATG	4020
109	GAGAGCGAGA	ACGAGAACAT	GCTGGCGGAG	TACGAGCAGC	GCCAGTACAC	CAGCGATTGC	4080
110	TGCAATAGCT	CCCGCGAGGG	CGACACCTGC	TCCTGCAGCG	AGGGATCCTG	TCTTACGCC	4140
111	GAGGCGGGCG	AGCCGGCGCC	TCGTCAAATG	ACTGCTAAGA	ACACCTAA		4188

113 (2) INFORMATION FOR SEQ ID NO: 2:

114 (i) SEQUENCE CHARACTERISTICS:

- 115 (A) LENGTH: 1395 amino acids
- 116 (B) TYPE: amino acid
- 117 (C) STRANDEDNESS: single
- 118 (D) TOPOLOGY: linear

119 (ii) MOLECULE TYPE: peptide

120 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

121	Met His Pro Met His Pro Glu Asn His Ala Ile Ala Arg Ser Thr Ser			
122	1	5	10	15
123	Thr Thr Asn Asn Pro Ser Arg Ser Arg Ser Ser Arg Met Trp Leu Leu			
124	20	25	30	
125	Pro Ala Trp Leu Leu Val Leu Val Ala Ser Asn Gly Leu Pro Ala			
126	35	40	45	
127	Val Arg Gly Gln Tyr Gln Ser Pro Arg Ile Ile Glu His Pro Thr Asp			
128	50	55	60	
129	Leu Val Val Lys Lys Asn Glu Pro Ala Thr Leu Asn Cys Lys Val Glu			
130	65	70	75	80
131	Gly Lys Pro Glu Pro Thr Ile Glu Trp Phe Lys Asp Gly Glu Pro Val			
132	85	90	95	
133	Ser Thr Asn Glu Lys Lys Ser His Arg Val Gln Phe Lys Asp Gly Ala			
134	100	105	110	
135	Leu Phe Phe Tyr Arg Thr Met Gln Gly Lys Lys Glu Gln Asp Gly Gly			
136	115	120	125	
137	Glu Tyr Trp Cys Val Ala Lys Asn Arg Val Gly Gln Ala Val Ser Arg			
138	130	135	140	
139	His Ala Ser Leu Gln Ile Ala Val Leu Arg Asp Asp Phe Arg Val Glu			
140	145	150	155	160
141	Pro Lys Asp Thr Arg Val Ala Lys Gly Glu Thr Ala Leu Leu Glu Cys			
142	165	170	175	
143	Gly Pro Pro Lys Gly Ile Pro Glu Pro Thr Leu Ile Trp Ile Lys Asp			
144	180	185	190	
145	Gly Val Pro Leu Asp Asp Leu Lys Ala Met Ser Phe Gly Ala Ser Ser			
146	195	200	205	
147	Arg Val Arg Ile Val Asp Gly Gly Asn Leu Leu Ile Ser Asn Val Glu			

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/971,172A

DATE: 07/25/2001

TIME: 15:27:30

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07252001\H971172A.raw

148	210	215	220
149	Pro Ile Asp Glu Gly Asn Tyr Lys Cys Ile Ala Gln Asn Leu Val Gly		
150	225	230	235
151	Thr Arg Glu Ser Ser Tyr Ala Lys Leu Ile Val Gln Val Lys Pro Tyr		240
152	245	250	255
153	Phe Met Lys Glu Pro Lys Asp Gln Val Met Leu Tyr Gly Gln Thr Ala		
154	260	265	270
155	Thr Phe His Cys Ser Val Gly Gly Asp Pro Pro Pro Lys Val Leu Trp		
156	275	280	285
157	Lys Lys Glu Glu Gly Asn Ile Pro Val Ser Arg Ala Arg Ile Leu His		
158	290	295	300
159	Asp Glu Lys Ser Leu Glu Ile Ser Asn Ile Thr Pro Thr Asp Glu Gly		
160	305	310	315
161	Thr Tyr Val Cys Glu Ala His Asn Asn Val Gly Gln Ile Ser Ala Arg		320
162	325	330	335
163	Ala Ser Leu Ile Val His Ala Pro Pro Asn Phe Thr Lys Arg Pro Ser		
164	340	345	350
165	Asn Lys Lys Val Gly Leu Asn Gly Val Val Gln Leu Pro Cys Met Ala		
166	355	360	365
167	Ser Gly Asn Pro Pro Pro Ser Val Phe Trp Thr Lys Glu Gly Val Ser		
168	370	375	380
169	Thr Leu Met Phe Pro Asn Ser Ser His Gly Arg Gln Tyr Val Ala Ala		
170	385	390	395
171	Asp Gly Thr Leu Gln Ile Thr Asp Val Arg Gln Glu Asp Glu Gly Tyr		
172	405	410	415
173	Tyr Val Cys Ser Ala Phe Ser Val Val Asp Ser Ser Thr Val Arg Val		
174	420	425	430
175	Phe Leu Gln Val Ser Ser Val Asp Glu Arg Pro Pro Pro Ile Ile Gln		
176	435	440	445
177	Ile Gly Pro Ala Asn Gln Thr Leu Pro Lys Gly Ser Val Ala Thr Leu		
178	450	455	460
179	Pro Cys Arg Ala Thr Gly Asn Pro Ser Pro Arg Ile Lys Trp Phe His		
180	465	470	475
181	Asp Gly His Ala Val Gln Ala Gly Asn Arg Tyr Ser Ile Ile Gln Gly		480
182	485	490	495
183	Ser Ser Leu Arg Val Asp Asp Leu Gln Leu Ser Asp Ser Gly Thr Tyr		
184	500	505	510
185	Thr Cys Thr Ala Ser Gly Glu Arg Gly Glu Thr Ser Trp Ala Ala Thr		
186	515	520	525
187	Leu Thr Val Glu Lys Pro Gly Ser Thr Ser Leu His Arg Ala Ala Asp		
188	530	535	540
189	Pro Ser Thr Tyr Pro Ala Pro Pro Gly Thr Pro Lys Val Leu Asn Val		
190	545	550	555
191	Ser Arg Thr Ser Ile Ser Leu Arg Trp Ala Lys Ser Gln Glu Lys Pro		560
192	565	570	575
193	Gly Ala Val Gly Pro Ile Ile Gly Tyr Thr Val Glu Tyr Phe Ser Pro		
194	580	585	590
195	Asp Leu Gln Thr Gly Trp Ile Val Ala Ala His Arg Val Gly Asp Thr		
196	595	600	605

RAW SEQUENCE LISTING

PATENT APPLICATION: US/08/971,172A

DATE: 07/25/2001

TIME: 15:27:30

Input Set : A:\seqlist.txt
 Output Set: N:\CRF3\07252001\H971172A.raw

197 Gln Val Thr Ile Ser Gly Leu Thr Pro Gly Thr Ser Tyr Val Phe Leu
 198 610 615 620
 199 Val Arg Ala Glu Asn Thr Gln Gly Ile Ser Val Pro Ser Gly Leu Ser
 200 625 630 635 640
 201 Asn Val Ile Lys Thr Ile Glu Ala Asp Phe Asp Ala Ala Ser Ala Asn
 202 645 650 655
 203 Asp Leu Ser Ala Ala Arg Thr Leu Leu Thr Gly Lys Ser Val Glu Leu
 204 660 665 670
 205 Ile Asp Ala Ser Ala Ile Asn Ala Ser Ala Val Arg Leu Glu Trp Met
 206 675 680 685
 207 Leu His Val Ser Ala Asp Glu Lys Tyr Val Glu Gly Leu Arg Ile His
 208 690 695 700
 209 Tyr Lys Asp Ala Ser Val Pro Ser Ala Gln Tyr His Ser Ile Thr Val
 210 705 710 715 720
 211 Met Asp Ala Ser Ala Glu Ser Phe Val Val Gly Asn Leu Lys Lys Tyr
 212 725 730 735
 213 Thr Lys Tyr Glu Phe Phe Leu Thr Pro Phe Glu Thr Ile Glu Gly
 214 740 745 750
 215 Gln Pro Ser Asn Ser Lys Thr Ala Leu Thr Tyr Glu Asp Val Pro Ser
 216 755 760 765
 217 Ala Pro Pro Asp Asn Ile Gln Ile Gly Met Tyr Asn Gln Thr Ala Gly
 218 770 775 780
 219 Trp Val Arg Trp Thr Pro Pro Pro Ser Gln His His Asn Gly Asn Leu
 220 785 790 795 800
 221 Tyr Gly Tyr Lys Ile Glu Val Ser Ala Gly Asn Thr Met Lys Val Leu
 222 805 810 815
 223 Ala Asn Met Thr Leu Asn Ala Thr Thr Ser Val Leu Leu Asn Asn
 224 820 825 830
 225 Leu Thr Thr Gly Ala Val Tyr Ser Val Arg Leu Asn Ser Phe Thr Lys
 226 835 840 845
 227 Ala Gly Asp Gly Pro Tyr Ser Lys Pro Ile Ser Leu Phe Met Asp Pro
 228 850 855 860
 229 Thr His His Val His Pro Pro Arg Ala His Pro Ser Gly Thr His Asp
 230 865 870 875 880
 231 Gly Arg His Glu Gly Gln Asp Leu Thr Tyr His Asn Asn Gly Asn Ile
 232 885 890 895
 233 Pro Pro Gly Asp Ile Asn Pro Thr Thr His Lys Lys Thr Thr Asp Tyr
 234 900 905 910
 235 Leu Ser Gly Pro Trp Leu Met Val Leu Val Cys Ile Val Leu Leu Val
 236 915 920 925
 237 Leu Val Ile Ser Ala Ala Ile Ser Met Val Tyr Phe Lys Arg Lys His
 238 930 935 940
 239 Gln Met Thr Lys Glu Leu Gly His Leu Ser Val Val Ser Asp Asn Glu
 240 945 950 955 960
 241 Ile Thr Ala Leu Asn Ile Asn Ser Lys Glu Ser Leu Trp Ile Asp His
 242 965 970 975
 243 His Arg Gly Trp Arg Thr Ala Asp Thr Asp Lys Asp Ser Gly Leu Ser
 244 980 985 990
 245 Glu Ser Lys Leu Leu Ser His Val Asn Ser Ser Gln Ser Asn Tyr Asn

VERIFICATION SUMMARY

PATENT APPLICATION: US/08/971,172A

DATE: 07/25/2001

TIME: 15:27:31

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\07252001\H971172A.raw

L:23 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]

L:24 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]

L:1196 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10

L:1210 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:10